AHFD

Aviation Human Factors Division
Institute of Aviation

University of Illinois
at Urbana-Champaign
1 Airport Road
Savoy, Illinois 61874

Display Formatting and Situation Awareness Model (DFSAM): An Approach to Aviation Display Design

Christopher D. Wickens

Technical Report AHFD-05-14/NASA-05-5

October 2005

Prepared for

NASA Ames Research Center
Moffett Field, CA

Contract NASA NAG 2-1535

NASA DFSAM Modeling Effort. October 24, 2005

Display Formatting and Situation Awareness Model (DFSAM): An Approach to Aviation Display Design

Christopher D. Wickens

Abstract

We describe a computational model, Display Formatting and Situation Awareness (DFSAM) that is designed to predict a figure of merit for an aviation display that has any of eight objectively quantifiable or definable features, such as its dimensionality (2D or 3D), clutter, or size. A set of 46 different experimental results (effect size estimates), and two meta analyses, extracted from 25 different studies conducted at University of Illinois, are analyzed and aggregated to estimate the net cost or benefit of the effect in question. These costs/benefits are referred to as **amalgamated performance units** (APUs). The APUs can then be combined for any particular display to predict its figure of merit, and the *difference* in figure of merit from any other display. The model is validated against the data from a high fidelity synthetic vision simulation, conducted with four different display formats, and was found to predict multitask flight control performance (r =0 .89) and traffic awareness response time (r = 0.81) and accuracy (r = 0.96). Constraints and limits of the model are discussed.

1. Overview and Modeling Approach

Various design guidelines exist for aviation display formatting. Often these exist as principles (Wickens, 2003, Reising Liggett and Munns, 1999), as well as “best practices” or Aerospace Recommended Practice (ARP) documents. What is often lacking are computational models that will “compute” how principles will aggregate or trade-off against each other, when certain designs incorporate a variety of such principles. A common example is how the benefits of close visual proximity between a display and the outside world view fostered by a HUD, will trade-off against the costs of clutter. While this modeling approach has been suggested in relatively static display interfaces (Tullis, 1988; Wickens Vincow, Schopper and Lincoln, 1997), it has been less evident in the dynamic world of aviation displays. At the same time, rapid developments in display technology have provided designers with greater freedom to format displays using a variety of innovative techniques, whose impact should be predicted.

The current report provides one example of such a computational modeling approach called *Display formatting situation awareness modeling,* or DFSAM. In the following we first describe the modeling approach, and extraction of parameters that characterize the influence of different format features. Then, in a validation, we apply this model to predict pilot performance differences between four different display formats in a synthetic vision system flight simulation.

More specifically, the goal of the DFSAM (Display Format Situation Awareness Model) model is to be able to compute a figure of merit of a particular display, on the basis of its features as that display is designed to perform a particular task:

The display features are those which are:

(a) elements that a display designer can alter (e.g., the location or size of the display within the cockpit)

(b) elements that have been used as independent variables in research carried out at Illinois (and funded by NASA), and have generated measures of pilot performance, thereby providing a basis for empirically validating the model.

The elements we consider are:

1. display size
2. display clutter
3. display overlay of separate data bases (e.g., traffic and weather)
4. display overlay of a head up display (HUD) (on the outside world)
5. degree of display separation
6. display highlighting/color coding
7. display prediction
8. display dimensionality (3D egocentric (immersed) vs. 2D coplanar

For each of these variables, we try to extract a measured “strength of effect”. For variables that are quantitative in form (e.g., display separation or size), this effect is a regression slope. For variables that are qualitative in form (e.g., overlay vs separate, 2D vs 3D), a single value measure is used.

One issue that has challenged our search for strength of effect measures, is amalgamating *statistical significance* (p value of effect) with *practical significance* (e.g., size of a difference in accuracy or response time). The former can be done independently of the nature of the study, but the latter must take into account the nature of the task (e.g. whether responses are measured in millisecond or seconds, or whether an effect is in accuracy or speed). Unless these differences are accounted for, one can have a situation in which a “marginally significant” 5 second cost for a particular display format, totally dominates highly significant _ second benefits of the same display from 2 two other studies. In this case, by summing the raw effects, the conclusion would be of a net 4 second cost (5- _ - _). Alternatively, considering only the role of statistical significance, the conclusion would be of a 2:1 (or greater) benefit for the display format in question. We resolve these ambiguities through amalgamation.

In order to amalgamate statistical significance (p value) and practical significance (raw data effect size), we have categorized the size of significant effects into two levels: small, (<10% of unit value) and large (>10% of unit value). We have also categorize “statistical significance” into three levels: (1) >.10, (2) >.10>p>.05, and (3) .05>p.

This allows us to create 5 levels of resolution of amalgamated effects:

		Effect Size	
		Small	Large
	>.10	0	0
P value	>10-.05	1	2
	<.05	2	3

In occasional circumstances, an effect was so overwhelmingly large, and highly significant, that it was allocated an effect size of 4. We can refer to these units as "**amalgamated performance units**", or **APU**s.

2. Quantification of APU Effects

In the following, a bracketed value (e.g., [2]) refers to the APU effect size estimate derived from the procedure above for a particular variable within a particular experiment. We have chosen to bold face any negative APUs. (e.g., **-[1.5]**). The "Study" is given a brief name identification, for which details can be found in the reference list.

2.1 Display Size

A series of experiments have examined the effects of display size, on tracking supported by that display, on the ability to make spatial judgments based on information within the display and on surveillance of hazards to be monitored on the display. In all of the findings below, the effect size is that which favors a large display (i.e., a positive APU indicates a large-display benefit). Because size is a quantitative variable, we characterize each study by the ratio of the large to the small display. The unit is of display radius (or width), rather than area.

Study	Size Ratio	Tracking	Spatial Judgments	Surveillance
Muthard E1	3:1	[3]	-	[0]
Muthard E2	3:1	[3]	-	[0]
Muthard E3	2:1	[2]	-	-
Muthard E4	3:1	[4]	-	-
Alex SVS2	1.5:1	[1]	**-[1]** (0 accuracy, small faster)	[3]
Kroft1	2:1		[0] (Large accurate Small faster)	
Kroft2	2:1		[2.5] (RT 15%, .01; acc: 10% .05)	

Conclusion. For tracking, the magnitude of the benefit closely tracks the ratio of the larger to the smaller display: that is, an average [2.6] benefit is observed for an average (across studies) 2.5 ratio of size. We can approximate this by assigning a benefit ratio of largeness = 1 X ratio of size. For spatial judgments, there is a mean large-display benefit of [1.5] across a mean size increase of 1.8. Thus the benefit is approximately 0.8 X ratio of size. Finally, for surveillance, a mean benefit of [1] is found for a mean size increase of 2.5. Thus the benefit is 0.4 X ratio of size.

2.2. Display Clutter

The effect of display clutter is somewhat challenging to quantify for several reasons. First local density clutter (extra irrelevant marks around an item to be encoded) exerts qualitatively different effects from global density clutter, which defines the total amount of marks on a display, both relevant and irrelevant (Wickens Vincow, Schopper and Lincoln, 1997). Second, clutter has very non-linear effects on display processing. Third, there is no firm consensus on how to measure clutter; so even though it is clearly an analog (rather than categorical) variable, it is not as easy to express in a regression equation as is, for example, the effect of size. The following three studies yielded clutter effects in which clutter cost is assigned a negative APU:

Study	Effect on Search Performance	
Muthard Expt 2	**-[1]**	
Kroft1 (unpublished)	**-[3]**	Global density
Nunes	**-[2]**	

Conclusion. We can therefore approximate a -[2] cost for a “more cluttered” display. However considerable caution remains because this measure does not account for the variety of ways in which clutter can be created, nor the different quantitative measures of clutter.

2.3 Display Overlay

The variable of display overlay is a categorical one, referring to whether two data bases overlay each other or are placed side-by-side. As such, it trades off the increased clutter of display overlay, against the increased scanning of a side-by-side array. The penalty of increased scanning is augmented to the extent that the two data bases contain spatially related information, so that they can be both expressed within the same coordinate system when overlaid, and that the information to be extracted from the two displays (whether overlaid or separate) pertains to the spatial relationship between them; that is, an **integration** task (Kroft and Wickens, 2003: Wickens and Carswell, 1995). In the following, the APU number portrays an **overlay advantage**, so a “separate advantage” is indicated by a negative sign. We distinguish between those tasks that require integration of the overlaid (or separated) data bases, and those that require the focus of attention on one or the other.

Study	Effect	Task Type	
	Integration	Focused	
O'Brien	[2]		CDTI * traffic conflict and weather display (NASA)
Helleberg	[1]		CDTI traffic and weather display (Rockwell Collins)
Kroft 1	[3]	**-[1]**	
Kroft 2	[2]	**-[2]** (-1 RT -3 accuracy if display is small)	

*CDTI refers to Cockpit Display of Traffic Information.

Conclusion. When the data need to be integrated, there is a [2] overlay benefit. When data must be processed independently, as in a focused attention search task in one domain only, there is a [1.5] overlay cost (or [1.5 separation benefit]).

2.4 HUD (head up display) Overlay

The HUD variable is closely related to the display overlay variable described in 2.3 and concerns the relative performance of tasks, as they are performed on a HUD, overlaying a display on some non-uniform background scene versus on a separate head down display. Such a contrast has little meaning if the out-the-window scene is blank (as in VMC). An important distinction can be made between tasks requiring focused attention on the display instruments, focused attention on the outside scene, and divided attention (or integration) between the two Wickens Ververs and Fadden, 2004).

The major source of input to this comparison is the results of the meta-analysis carried out by Fadden et al (1999) of all such HUD studies, and therefore including both NASA Illinois studies, as well as all other HUD studies, in which the format of a HUD and head down presentation is otherwise equivalent. Positive signs on the APU measures are those that favor the HUD.

Study	Focused Near (Display)	Focused Far (Outside)	Divided (Integration)
Fadden meta-analysis	[2]	**-[2]** (cruise flight tracking. No outside events)	[3]
		-[1] (response to off-normal)	
Fadden-Ververs			
Tracking	-	[0]	[2] (taxi) **-[2]** (final approach)
Event detection	[2]	[3] [0] (response to off normal event)	

Conclusion. In calculating the net APU, we have given twice the weighting to effects revealed by the Fadden et al meta-analysis, as to effects revealed by the individual study (which did not enter into the meta-analysis), given the considerably larger sample size of studies in the former.

Focused attention to display information:	HUD benefit = [2.0]
Focused attention on far domain: normal events	HUD cost = **-[0.5]**
Focused attention on far domain: off-normal (unexpected)	HUD cost = **-[1.0]**
Divided attention, tracking:	HUD benefit: = [1.5]

Therefore the HUD appears to have its greatest benefit on processing displayed information, a benefit that generally dominates the smaller cost to processing less-visible far domain events.

2.5 Display Separation

The display separation variable describes the effect of the degree of separation between two already separated displays (e.g., it does not include separation = 0, which characterizes the overlay condition in either 2.3 or 2.4 above). As such it represents the penalty of scanning between two displays, as a function of the amount of separation. We have tried to convert the separation variable into degrees of visual angle between the mid-points of the relevant displays.

Three studies were located that systematically varied the separation between displays. The range of separations examined is shown:
1. Schons and Wickens (5 degrees to 23 degrees vertically),
2. Wickens Dixon and Seppelt (8 degrees to 45 degrees horizontally) and
3. Horrey Wickens and Consalus (19 degrees to 37 degrees vertically).

All studies had a tracking task located at the top or left and a discrete task or another tracking task located at varying separations below or to the right. Because Wickens Dixon and Seppelt only measured tracking error while the concurrent task was performed, while the other two studies measured tracking error during the entire trial (approximately 40% of which was occupied by concurrent task processing), the tracking error cost of Wickens Dixon and Seppelt was converted to a "diluted" measure of the tracking error that would have been observed had it been averaged across the entire period (e.g., including those intervals of time when a visual event was not being processed).

Tracking error scores are expressed as a percent RMS (root mean squared) error increase per unit visual angle separation. We have also employed a common metric of % increase/10 degrees across all studies for comparison purposes. RT measures are not converted in this way because of widely varying (across the three studies) base rates of the RT measure (depending on the nature of the discrete task). Instead, the latter are expressed as the decrement size (0-4) using our standard APU decrement scale: [X].

	Tracking	Discrete task RT
1. Schons & Wickens	8%/20 degrees = **4%/10 deg** (diluted)	[1]
2. Wickens Dixon et al	80%/37 deg (full time) = 30%/37 deg (diluted) **8%/10 deg** (diluted)	[2]
3. Horrey & Wickens	0%/20 deg = **0%/10 deg** (diluted)	[2]

Conclusion. There is some disparity across these estimates; particularly the absence of cost in Horrey Wickens and Consalus is somewhat puzzling. It may represent the particular nature of the tracking task, in this case taxiing a simulated aircraft, such that tracking error was robust to short glances away (downward) from the outside world where the taxi way was seen. We can pool the three studies to estimate an error increase cost of 4%/10 degrees. In order to convert this back to an APU cost, if we assume that a 10% cost is at the borderline between [1] and [2] (ie., 10% cost = [1.5]), then the 4% cost is associated with a [1]/ 10 deg.

For the discrete task we can estimate a mean penalty of [1.5] for 18 degrees of eccentricity (study 1 and 3), and [2] for 37 degrees of increasing eccentricity (study 2). Taking a weighted average across the three studies after first converting study 2 to a [1]/18 degrees yields a penalty of [1.3]/18 degrees or, converting it to the common unit of penalty/10 degrees, this becomes 0.7/10 degrees. One important point in concluding this section is that costs on the two tasks will be modulated by priorities (Horrey et al). That is, high priority given to the discrete task will impose a relatively greater cost of separation on the tracking task.

2.6. Display Highlighting/Color Coding

The display highlighting variable describes the benefits gained by rendering certain parts of a display, relevant to the task at hand, more intense or "highlighted" than background elements. In the current section, this also incorporates the use of color coding. Generally the effectiveness of highlighting will be to reduce the amount of material that needs to be visually searched in order to find elements that are the focus of the task (e.g., highlighting all aircraft at the same altitude as ownship). Hence most of our evaluations are of the benefits for a task requiring the focus of attention on the highlighted class. However sometimes there may be benefits to highlighting for tasks that depend on both highlighted and non-highlighted items (ie., divided attention). This is because highlighting half the items can better organize the space. We consider these two classes separately. Positive values indicate highlighting benefits.

Study	Focused Attention	Integration
Podczwerinski	[1]	
Kroft 2	[0]	
Wickens &Martens	[0] [3]	[1] [2]
Muthard/Wickens	[0]	
Nunes	[3]	

Conclusion. Collectively focused attention benefits of highlighting are [1.2], and integration benefits are [1.5].

2.7. Prediction

Prediction refers to display enhancements that project future control requirements. In most tracking or flight control studies these can be differentiated into future flight path requirements: called *preview* (such as the flight path highway in the sky), and future aircraft predicted position, called *prediction*, such as a flight path vector. We separate the benefits of each below. Four studies explicitly examined the presence or absence of one or both of these elements. One study (Doherty) also presented a meta-analysis of the collective results of several others. As with our treatment of HUDs above (section 2.4), the output from this analysis is given twice the weight of each of the other individual studies by themselves when the collective results are analyzed. Where relevant, specific benefits of preview are distinguished by those to lateral tracking and to vertical tracking.

Study	Preview (Pathway) Benefit	Prediction Benefit
Doherty meta-analysis	[3]	[3]
Doherty study	[3] lat [3] vert	[0] lat [2] vert
Alex SVS1	[0] lat [2] vert	-
Morphew (CDTI-traffic projection)	-	[3]
Fadden-Ververs	[3] -	[3]

Conclusion. Prediction: 11/5 = mean prediction benefit of [2.2]. Preview: 14/6 = mean preview benefit of [2.3]. It should be noted that there are some interactions present in the data, not represented in the numbers. That is, there is evidence from Doherty, that the benefits of prediction are greater when preview is present, than when it is not.

2.8. Dimensionality

A great number of studies supported by NASA have compared 2D with 3D displays. In all of these, "2D" has actually consisted of a suite of two "co-planar" displays, so that identical depth and distance information is conveyed as with the 3D display. Furthermore, across different studies, the 3D display has adopted two qualitatively different viewpoints, an exocentric viewpoint, where the pilot's own ship position is visible in the display, and an egocentric display, in which the aircraft's position corresponds to the viewpoint of the display. Finally, studies have examined both detection/decision making (assessed by RT and accuracy) and tracking/flight control (assessed by RMS error).

A. 2D Co-planar vs 3D exocentric. In the following we present a "coplanar advantage", so a negative sign indicates a 3D advantage.

Study		Detection/Decision	Tracking
Alwick SVS4	E1	[2] (vertical)	
		[0] (lateral)	
	E2	[1] (vertical)	
		[0] lateral	
Thomas CDTI			
	E1, E2, E3, RT and Accuracy: no difference: (i.e., [0,0,0,0,0,0])		
Alwickmer CDTI			
	E1	[2]	
	E2	[0]	
	E3	[0]	
Muthard		[1]	
Olmos Wick Chudy			[0]
		-[2] (RT)	
		[3] Accuracy	
Olmos Liang			**-[1]** lat
			-[1] vert
		[0]	
		[0]	
WickLiangPrevett E1			[0]
			[3]
		[0]	
		[3]	
		-[3]	
		-[3]	
			[1]
			[0]
WickensLiangPrev E2		[1]	
		[0]	
Wickens & Prevett			[2]
			[1]
			-[1]

B. Co-planar vs Egocentric. Here we represent the 3D egocentric advantage, with a positive value.

	Tracking	
Wickens & Prevett	[3]	
	[3]	
	[2]	
	[0]	
Haskell & Wickens	[3]	
	[3]	
	-[1]	(airspeed tracking)

C. 3D Egocentric vs Exocentric Here again, we represent an egocentric advantage.

Wickens & Prevett	[3]
	[3]
	-[1]
Doherty	[3]
	[3]

Conclusion.

Coplanar vs Exocentric. Detection/diagnosis: [0.25].

Tracking: [0.50] (both indicating a small coplanar benefit), suggesting that the costs of 3D ambiguity slightly dominate the costs of scanning.

With egocentric displays (tracking only) there is a substantial "immersed benefit"
3D egocentric vs Coplanar [2]
3D egocentric vs. 3D Exocentric [2.2]

3. Validation Experiment

The weightings described in section 2 can be combined in a linear additive model to predict a figure of merit for a particular display that possesses the different features (e.g., an overlaid 3D exocentric display). In order to validate this model, we compared its predictions, against the performance data from an experiment *that was intentionally withheld from those generating the weightings.*

In this validation experiment (Wickens, Alexander, Thomas Horrey Nunes and Hardy, 2004), four different display layouts were contrasted in a synthetic vision system (SVS) evaluation. These four, shown in Figure 3.1, were defined by the 2X2 combination of (a) presence or absence of a 3D flight path tunnel in the sky (prediction), and (b) whether or not the flight instruments overlaid the SVS terrain, traffic and guidance display. (Note that these four conditions, flown either in IMC or VMC had also been used to validate the A-SA scanning model, as reported in Wickens McCarley Thomas Horrey Alexander and Zheng, 2005). We label these four conditions: overlay tunnel, overlay/data link (or just "overlay"), separate tunnel and separate datalink (or just “separate”). In variable (a), when the 3D preview was absent, pilots needed to assess command flight path information from separated co-planar displays. In variable (b), when the flight instruments were separated, pilots needed to scan 15 degrees from the center of the SVS display to the center of the instrument cluster.

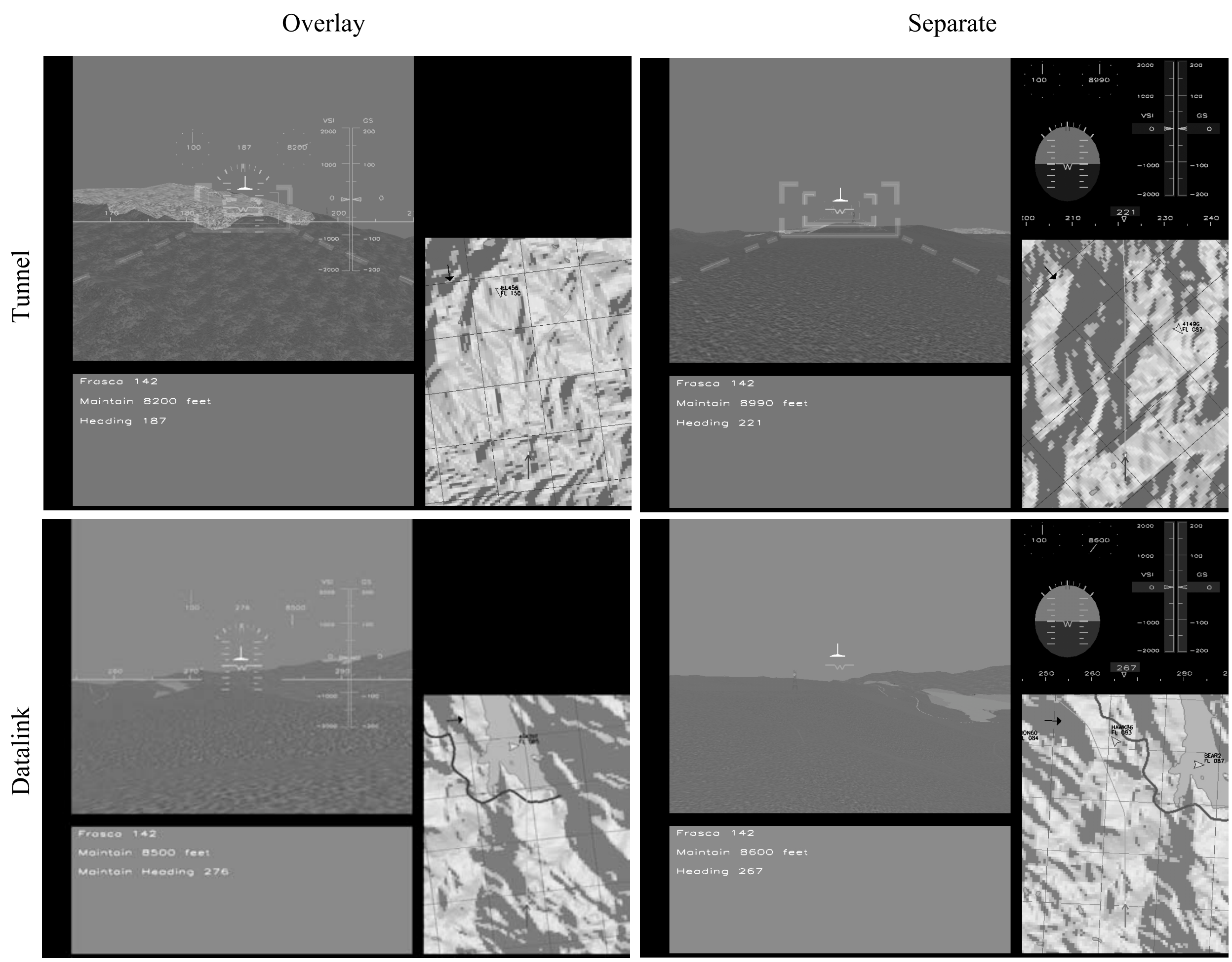

Figure 3.1. The Four Display layouts used for DFSAM validation.

Pilots were assigned three tasks, two relevant here for model validation: maintain the lateral and vertical flight path, and report traffic as it appeared on the 3D SVS panel. A third task, reporting traffic changes on the navigational display was not employed for validation, since the information representing this task was unaffected by any of the four display layouts. Importantly however, the task contributed to the overall cockpit task workload. Table 3.1 shows the APU benefits or penalties associated with the different conditions, for the two different tasks, as these values were derived from the analysis in section 2. That is, for example, in the top row of the table, we can find the benefit of overlay to tracking [2], and its cost to detection (-[1.5]) as revealed in section 2.3. The net influence of all components for a particular task is shown in bold face within Table 3.1.

Table 3.1. Net APUs for the four different conditions and two different tasks.

Condition	Tracking	Traffic Detection
Overlay (OL) Tunnel	OL [2] + 3D [2] + Preview [2.5] = **[6.5]**	OL = **-[1.5]**
Overlay	OL = **[2.0]**	OL = -**[1.5]**
Separated Tunnel Distance Penalty:	0.6/10 deg X 15 deg =-[0.9] + 3D ego [2] + preview [2.5] = **[3.6]**	0.7/10 degX15 deg = **-[1.0]**
Separated Distance =	**-[0.9]**	**= -[1.0]**

Given this derivation, these predicted performance levels were correlated against obtained measures of Tracking RMS vertical and lateral deviations, and Traffic detection RT and error rate. These values are shown in Table 3.2, along with the predicted values for each variable as carried down from Table 3.1.

Table 3.2. Predicted [ ] and obtained performance values four the four display formats.

Condition		Tracking			Traffic Detection	
		Lateral	Vertical		RT (sec)	Error Rate
OL Tunnel	**[6.5]**	7.5	5.2	**-[1.5]**	17.0	0.16
OL	**[2.0]**	71.1	25.5	**-[1.5]**	20.6	0.18
Separate Tunnel	**[3.6]**	8.2	5.7	**-[1.0]**	9.6	0.10
Separate	**-[0.9]**	81.5	31.6	**-[1.0]**	15.1	0.07

On the basis of these values, four different product moment correlation values were computed, correlating predicted versus obtained measures of lateral and vertical tracking and traffic detection RT and error rate. These values are:

Lateral tracking	r = -0.88
Vertical Tracking:	r = -0.90
Traffic Response time:	r = -0.81
Traffic Error rate:	r = -0.96

The scatter plots generating these correlations are shown in Figure 3.2. The negative sign indicates that higher predicted APUs are associated with less error and shorter response times.

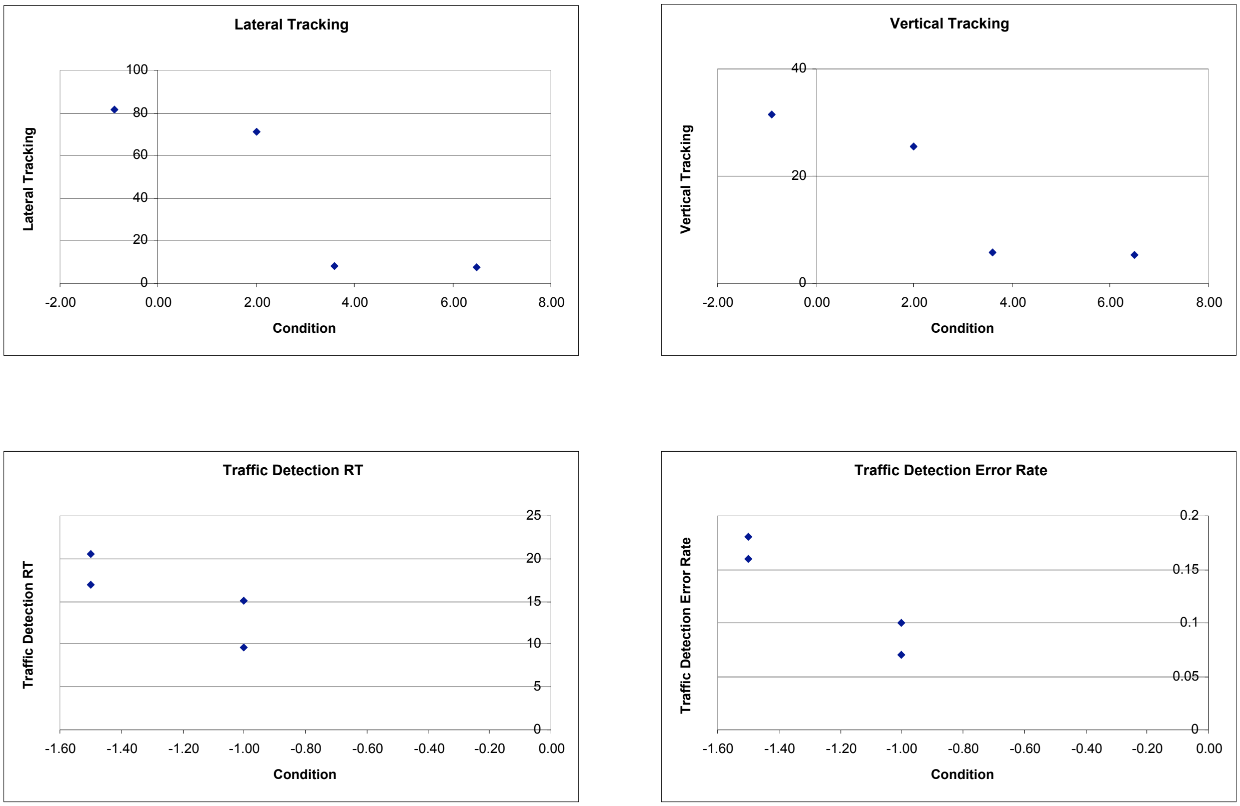

Figure 3.2. Scatter plots representing the DFSAM model validation.

While these correlations are all strong, in supporting model predictions, examination of the scatter plots reveals possible model adjustments to increase the prediction quality still further. In particular, it appears that in three of the four correlations, the benefit of the 3D preview is underestimated. For both tracking tasks (top two graphs), were the two sets of points on the left, farther separated from the two on the right (a manifestation of a greater reward for the tunnel display), the linear correlation would be greater. For the traffic detection RT measure there was no “reward” added for using the tunnel display, and hence the tunnel and non-tunnel predictions are of the same value. However if, (as revealed by secondary task measures in other studies), it is assumed that the resource demands of tracking are reduced by the tunnel, and hence a “workload penalty” added to the non-tunnel conditions would shift the two upper points to the left, this too would increase the model fit. Such a shift would not however benefit the error rate correlation.

While a workload penalty component might therefore add to the overall predictability of the model, such an approach was not considered here, because workload is a property that requires empirical measurement, whereas our display layout model is based entirely on a-priori measurable quantities.

It should be noted that one fundamental assumption was made here, that substantially influenced the results. Specifically, we chose the overlay values (section 2.3), rather than the HUD overlay values (section 2.4) to characterize the displays on the left side of Figure 3.1. Our rationale for doing so was that what we evaluated was not a true HUD, viewed over the outside scene, and that the "separated" condition was not characterized by accommodation differences, and downward viewing, characteristic of most HUD research. However, had we used the values in section 2.4, the overlay penalty to focused attention on the far domain would have been reduced to [0.5]. In the lower two scatter plots of Figure 3.2, this would have shifted the two values on the left of each plot from [1.5] to [0.5], and reverse the correlations.

However, we also note that in deriving the overlay penalty values in section 2.4 (in particular, the abovementioned [0.5] value), we provided a 2:1 ratio of weighting the meta-analysis (which showed a –[2] cost) to the single study reported (which showed a [3] benefit). Here a strong case could be made that this weighting ratio, should have instead been more like 30:1, reflecting the number of studies entering into the meta-analysis. Had this been done, a penalty approaching the [2.0] value would have been computed, actually strengthening the correlation.

4. Conclusions

In conclusion, we note the relative success of the DFSAM model in making predictions across a small number of independently collected data points. However it is also legitimate to address additional issues about the validation effort, two of which we address below.

First, is the flight control measure, which contributes to half of the validation, truly a "situation awareness" measure, representative of a model that purports to be based on situation awareness? It is true that flight control depends on many factors above and beyond the perceptual-cognitive ones associated with situation awareness. However given that our assessment of flight control in the current simulation was done in the context of two competing visual tasks (outside, and Nav display traffic detection), this multi-task context forces imposes on the pilot an attentional sampling routine, which is a key element underlying stage 1 situation awareness in Endsley's model. Indeed we note here the validation of a separate SA scan model based on the same flight control measures described here (Wickens et al, 2005). Furthermore, in any flight control task, whether single or dual task, the more extensive predictive requirements, imposed by higher order flight control dynamics, by definition, impose on Endsley's stage 3 SA.

The second issue regards the absence of two important display variables that were additionally considered in the validation effort: display modality and automation reliability. We eventually decided to exclude modality (the choice to display information auditorally rather than visually) after examining the impact of a large number of studies that had compared auditory with visual presentation (e.g., Wickens Dixon and Seppelt, 2001; Wickens Goh Helleberg Horrey and Talleur, 2003). The collective results of these studies were extremely variable, not allowing us to extract a single estimate APU with much confidence. It also turns out that a large source of this variability is the visual separation that characterizes the two visual channels whose task performance is contrasted with that when one of them is auditory. Variance attributable to this property **is** well captured by our section 2.5. Finally, we find that when visual-visual separation is quite small, the APU value appears to be about [0], suggesting that there might be little value added to including this measure.

With regard to automation reliability, while this variable has factored into a good bit of our research (e.g., Wickens Gempler and Morphew, 2000), we have chosen not to include it here because (a) it is not truly a display property, and (b) its effects lend itself to a different form of computational modeling that can be found in Wickens and Dixon (2005).

5. References

Key to Ambiguous References.

In the following list, any reference abbreviations in the tables in sections 2.1-2.8 that may be ambiguous, are attached to unambiguous labels that can be found in the reference list below. The references for all other labels in the tables can easily be unambiguously found in the reference list.

Muthard E1 – E4. Muthard and Wickens (2005) experiments 1-4
Alex SVS1: Alexander and Wickens (2005) Expt 1
Alex SVS2: Alexander and Wickens (2005) Expt 2
Kroft 1: Paul Kroft. Unpublished manuscript. University of Illinois
Kroft 2: Kroft and Wickens, 2001, 2003
Fadden Met-analysis. Fadden Ververs and Wickens, 1998
Fadden-Ververs: Fadden Ververs and Wickens, 2001.
Morphew: Wickens Gempler and Morphew, 2000
AlwickSVS4: Alexander and Wickens, 2005.
Alwickmer: Alexander Wickens and Morphew, 2005.

Alexander, A. L., & Wickens, C. D. (2005). 3D n*avigation and integrated hazard display in advanced avionics: performance, situation awareness, and workload* (AHFD-05-10/NASA-05-2). Savoy, IL: University of Illinois, Aviation Human Factors Division.

Alexander, A., Wickens, C.D., & Hardy, T. J. (in press). Synthetic vision systems: The effects of guidance symbology, display size, and field of view. *Human Factors*.

Alexander, A. L., Wickens, C.D., & Merwin, D. (2005). Perspective and Co-planar cockpit display of traffic information: implications for maneuver choice, flight safety and mental workload. *International Journal of Aviation Psychology, 15*(1), 1-21.

Doherty, S. M., & Wickens, C. D. (2000). *An analysis of the immersed perspective flight path display benefit: Benefit of preview, prediction, and frame of reference* (ARL-00-5/NASA-00-1). Savoy, IL: University of Illinois, Aviation Res. Lab.

Fadden, S., Ververs, P., & Wickens, C.D. (1998). Costs and benefits of head-up display use: A meta-analytic approach. *Proceedings of the 42nd Annual Meeting of the Human Factors & Ergonomics Society*. Santa Monica, CA: Human Factors Society.

Fadden, S., Ververs, P.M., & Wickens, C.D. (2001). Pathway HUDS: Are they viable? *Human Factors*, *43*(2), 173-193.

Haskell, I.D., & Wickens, C.D. (1993). Two- and three-dimensional displays for aviation: A theoretical and empirical comparison. *International Journal of Aviation Psychology, 3*(2), 87-109.

Horrey, W.J., Wickens, C.D., & Consalus, K.P. (2005). *The distracted driver: Modeling the impact of information bandwidth, in-vehicle task priority, and spatial-separation on driver performance and attention allocation* (AHFD-01-11/GM-05-02). Savoy, IL: University of Illinois, Aviation Human Factors Division.

Kroft, P., & Wickens, C.D. (2001). *The display of multiple geographical data bases: Implications of visual attention* (ARL-01-2/NASA-01-2). Savoy, IL: University of Illinois Aviation Res. Lab.

Kroft, P., & Wickens, C.D. (2003). Displaying multi-domain graphical database information: An evaluation of scanning, clutter, display size, and user interactivity. *Information Design Journal, 11*(1), 44-52.

Muthard, E. K., & Wickens, C. D. (2002). *Factors that mediate flight plan monitoring and errors in plan revision: An examination of planning under automated conditions* (AHFD-02-11/NASA-02-8). Savoy, IL: University of Illinois, Aviation Human Factors Division.

Muthard, E.K. & Wickens, C.D. (2005). *Display size contamination of attentional and spatial tasks: An evaluation of display minification and axis compression* (AHFD-01-12/NASA-05-03). Savoy, IL: University of Illinois, Aviation Human Factors Division.

Muthard, E. & Wickens, C.D. (2005, in press). Pilots strategically compensate for display enlargements in surveillance and flight control tasks. *Human Factors.*

O'Brien J.V., & Wickens, C.D. (1997). *Cockpit displays of traffic and weather information: Effects of dimension and data base integration* (ARL-97-3/NASA-97-1). Savoy, IL: University of Illinois Aviation Res. Lab.

Olmos, O., Liang, C-C., & Wickens, C.D. (1997). Electronic map evaluation in simulated visual meteorological conditions. *International Journal of Aviation Psychology, 7*(1), 37-66.

Olmos, O., Wickens, C.D., & Chudy, A. (2000). Tactical displays for combat awareness: An examination of dimensionality and frame of reference concepts and the application of cognitive engineering *The International Journal of Aviation Psychology, 10*(3), 247-271.

Nunes, A., & Wickens. C.D. (2005) Clutter effects in CDTI. Unpublished manuscript

Podczerwinski, E., Wickens, C.D., & Alexander, A.L. (2002). *Exploring the "out of sight, out of mind" phenomenon in dynamic settings across electronic map displays* (ARL-01-8/NASA-01-4). Savoy, IL: University of Illinois Aviation Res. Lab.

Reising, J., Liggett, K, & Munns, R. (1999). Controls, Displays, and Workplace Design. In D. Garland, J., Wise and V.D. Hopkin (Eds). *Handbook of Aviation Human Factors*. Mahwah, N.J.: Erlbaum.

Schons, V.W., & Wickens, C.D. (1993). *Visual separation and information access in aircraft display layout* (ARL-93-7/NASA-A3I-93-1). Savoy, IL: University of Illinois Aviation Res. Lab.

Thomas, L. C., & Wickens, C. D. (2005). *Effects of display dimensionality, conflict geometry, and time pressure on conflict detection and resolution performance using a cockpit display of traffic information* (AHFD-05-4/NASA-05-1). Savoy, IL: University of Illinois, Aviation Human Factors Division.

Wickens, C. D., Alexander, A. L. & Hardy, T. J. (2003). *The primary flight display and its pathway guidance: Workload, performance, and situation awareness* (Final Technical Report AHFD-03-02/NASA-03-01). Savoy, IL: University of Illinois, Aviation Human Factors Division.

Wickens, C.D., Dixon, S.R., & Seppelt, B. (2002) *In vehicle displays and control task interference: the effects of display location and modality* (AHFD-02-7/NASA-0205). Savoy, IL: University of Illinois, Aviation Human Factors Division.

Wickens, C.D., Gempler, K., & Morphew, E. (2000)/ Workload and reliability of predictor displays in aircraft traffic avoidance. *Transportation Human Factors. 2(2). 99-126.*

Wickens, C.D., & Helleberg, J. (1999). *Interactive perspective displays for airborne hazard awareness* (Final Technical Report ARL-99-1/ROCKWELL-99-1). Savoy, IL: University of Illinois Aviation Res. Lab.

Wickens, C.D., Liang, C-C, Prevett, T., & Olmos, O. (1996). Electronic maps for terminal area navigation: Effects of frame of reference and dimensionality. *International Journal of Aviation Psychology, 6*(3), 241-271.

Wickens, C.D., Martens, M., Alexander, A., & Ambinder, M. (2004). The role of highlighting in visual search through maps. *Spatial Vision, 1-16.*

Wickens, C. D., McCarley, J. S., Alexander, A. L., Thomas, L. C., Ambinder, M. & Zheng, S. (2004). *Attention-situation awareness (a-sa) model of pilot error* (AHFD-04-15/NASA-04-5). Savoy, IL: University of Illinois, Aviation Human Factors Division.

Wickens, C.D., & Prevett, T. (1995). Exploring the dimensions of egocentricity in aircraft navigation displays. *Journal of Experimental Psychology: Applied, 1*(2), 110-135.

Wickens, C.D., Vincow, M., Schopper, R. & Lincoln, J. (1997) Computational models of human performance in the design and layout of controls and displays. *SCERIAC SOAR Report 97-22.* Wright Patterson Air Force Base.